Supplementary Material for "Online Estimation for Functional Data"

In this Supplementary Material, we first derive the asymptotic distributions of the proposed mean and covariance estimates, i.e., Theorem 1 and 2, in Section S.1 and S.2, respectively. The convergence of bandwidth selection and lower bound of relative efficiencies are shown in Section S.3 and S.4.

S.1. PROOF OF THEOREM 1

Proof. This proof is in analogy to the proof of Theorem 3.1 in Zhang and Wang (2016). Denote the current time by K and recall $\left\{\widetilde{\eta}_{\mu,k}^{(K)} :, k = 1, \ldots, K\right\}$ is the pseudo-bandwidth chain at time K. We omit the superscript "(K)" of $\widetilde{\eta}_{\mu,k}^{(K)}$, $\widetilde{h}_{\mu}^{(K)}$ in (18), $\rho_{\mu,j}^{(K)}$ in (14) and $\widetilde{\mu}^{(K)}$ in (10) in the proof when no confusion exits. The online estimate can be written as

$$\widetilde{\mu}(t) = \frac{R_0 Q_2 - R_1 Q_1}{Q_0 Q_2 - Q_1^2}$$

where for $r = 0, 1, 2$,

$$Q_r = \frac{1}{S_{K,1}} \sum_{k=1}^{K} \sum_{i=1}^{n_k} \sum_{j=1}^{m_{ki}} W_{\widetilde{\eta}_{\mu,k}}(T_{kij} - t) \left(\frac{T_{kij} - t}{\widetilde{\eta}_{\mu,k}}\right)^r = \sum_k \omega_k Q_{r,k},$$

$$R_r = \frac{1}{S_{K,1}} \sum_{k=1}^{K} \sum_{i=1}^{n_k} \sum_{j=1}^{m_{ki}} W_{\widetilde{\eta}_{\mu,k}}(T_{kij} - t) \left(\frac{T_{kij} - t}{\widetilde{\eta}_{\mu,k}}\right)^r Y_{kij} = \sum_k \omega_k R_{r,k}.$$

We further denote for $r = 0, 1$,

$$\widetilde{G}_r = \frac{1}{S_{K,1}} \sum_{k=1}^{K} \sum_{i=1}^{n_k} \sum_{j=1}^{m_{ki}} W_{\widetilde{\eta}_{\mu,k}}(T_{kij} - t) T_{kij}^r = \sum_k \omega_k \widetilde{G}_{r,k},$$

and define

$$\widetilde{\mu}'(t) = \frac{1}{\widetilde{\eta}_{\mu,k}} \frac{R_1 - R_0 Q_1/Q_0}{Q_2 - Q_1^2/Q_0}.$$

Note that

$$Q_1 = \sum_k \omega_k Q_{1,k} = \sum_k \omega_k \frac{\widetilde{G}_{1,k} - t\widetilde{G}_{0,k}}{\widetilde{\eta}_{\mu,k}},$$

one can derive

$$\widetilde{\mu}(t) = \frac{R_0}{Q_0} - \sum_k \omega_k \frac{\widetilde{G}_{1,k}}{Q_0} \widetilde{\mu}_k'(t) + t \sum_k \omega_k \frac{\widetilde{G}_{0,k}}{Q_0} \widetilde{\mu}_k'(t).$$

We also define

$$\overline{\mu}(t) = \frac{R_0}{Q_0} - \sum_k \omega_k \frac{\widetilde{G}_{1,k}}{Q_0}\mu'(t) + t\sum_k \omega_k \frac{\widetilde{G}_{0,k}}{Q_0}\mu'(t), \tag{S.1}$$

where $\mu'(\cdot)$ is the first derivative of $\mu(\cdot)$. Then

$$\widetilde{\mu}(t) - \overline{\mu}(t) = -\sum_k \omega_k \frac{Q_{1,k}}{Q_0} \cdot \frac{Q_0(R_1 - \mu Q_1 - \widetilde{\eta}_{\mu,k}\mu' Q_2) - Q_1(R_0 - \mu Q_0 - \widetilde{\eta}_{\mu,k}\mu' Q_1)}{Q_2 Q_0 - Q_1^2}. \tag{S.2}$$

It is straightforward to show that both Q_0 and $Q_0Q_2 - Q_1^2$ are positive and bounded away from 0 with probability tending to one,

$$Q_1 = O_p\left\{\rho_{\mu,1} + \left(\frac{\rho_{\mu,-1}}{S_{K,1}}\right)^{\frac{1}{2}}\right\},$$

$$(R_1 - \mu Q_1 - \widetilde{\eta}_k \mu' Q_2), (R_0 - \mu Q_0 - \widetilde{\eta}_k \mu' Q_1) = O_p\left(\rho_{\mu,2} + \rho_{\mu,1}\widetilde{\eta}_{\mu,k} + \left\{\frac{\rho_{\mu,-1}}{S_{K,1}}\right\}^{\frac{1}{2}} + \left\{\frac{S_{K,2}}{S_{K,1}^2}\right\}^{\frac{1}{2}}\right).$$

Then by the Cramér-Wold device and Lyapunov condition, we can achieve the asymptotic joint normality of $(R_0 - \mathrm{E}R_0, \widetilde{G}_1 - \mathrm{E}\widetilde{G}_1, \widetilde{G}_0 - \mathrm{E}\widetilde{G}_0)$, where the rate of convergence is $\min[(\rho_{\mu,-1}/S_{K,1})^{1/2}, \{Q_2/S_{K,1}^2\}^{1/2}]$.

Explicitly for $r, r' = 0, 1$,

$$\mathrm{E}(\widetilde{G}_r) = t^r f(t) + \frac{1}{2}\alpha(W)\{2rf'(t) + t^r f''(t)\}\rho_{\mu,2} + o_p(\rho_{\mu,2}),$$

$$\mathrm{E}(R_0) = \mu(t)f(t) + \frac{1}{2}\alpha(W)\{\mu(t)f''(t) + 2\mu'(t)f'(t) + \mu'' f(t)\}\rho_{\mu,2} + o_p(\rho_{\mu,2}),$$

$$\mathrm{Cov}(\widetilde{G}_r, \widetilde{G}_{r'}) = \left(\frac{\rho_{\mu,-1}}{S_{K,1}}\right) R(W) t^{r+r'} f(t) + o_p\left(\frac{\rho_{\mu,-1}}{S_{K,1}}\right),$$

$$\mathrm{Var}(R_0) = \left(\frac{\rho_{\mu,-1}}{S_{K,1}}\right) R(W)(\mu^2(t) + \gamma(t,t) + \sigma^2) f(t) + \left(\frac{S_{K,2}}{S_{K,1}^2}\right)\gamma(t,t)f(t)^2 + o_p\left(\frac{\rho_{\mu,-1}}{S_{K,1}} + \frac{S_{K,2}}{S_{K,1}^2}\right),$$

$$\mathrm{Cov}(R_0, \widetilde{G}_r) = \left(\frac{\rho_{\mu,-1}}{S_{K,1}}\right) R(W) t^r \mu(t) f(t) + o_p\left(\frac{\rho_{\mu,-1}}{S_{K,1}}\right).$$

From the delta method, $\overline{\mu}(t)$ follows the asymptotic normality that

$$\Gamma_\mu^{-1/2}\left\{\overline{\mu}(t) - \mu(t) - \frac{1}{2}\alpha(W)\mu^{(2)}(t)\rho_{\mu,2} + o_p(\rho_{\mu,2})\right\} \xrightarrow{d} N(0,1),$$

where

$$\Gamma_\mu = R(W)\frac{\gamma(t,t) + \sigma^2}{S_{K,1}f(t)}\rho_{\mu,-1} + \gamma(t,t)\frac{S_{K,2}}{S_{K,1}^2}.$$

When $\overline{m}_{\mu,K}/(N_K)^{1/4} \to 0$ and $\widetilde{h}_\mu \asymp S_{K,1}^{-1/5}$, the first term in Γ_μ above dominates. When $\overline{m}_{\mu,K}/(N_K)^{1/4} \to C$ and $\widetilde{h}_\mu \asymp S_{K,1}^{-1/5}$, where $0 < C < \infty$, the two terms of Γ_μ are of the same order as $\widetilde{h}_\mu^4$. When $\overline{m}_{\mu,K}/(N_K)^{1/4} \to \infty$ and $\widetilde{h}_\mu \asymp S_{K,1}^{-1/5}$, the second term of Γ_μ dominates and the bias becomes negligible. Noting (S.2), the proof of Theorem 1 is completed. ∎

S.2. PROOF OF THEOREM 2 AND LEMMA 1

We omit the superscript "(K)" of $\widetilde{\eta}_{\gamma,k}^{(K)}$, $\widetilde{h}_\gamma^{(K)}$, $\rho_{\gamma,j}^{(K)}$and $\widetilde{\gamma}^{(K)}$ in this section when no confusion exits. The proof of Lemma 1 is similar to the proof of Theorem 2, with $\rho_{\gamma,j}, j \in \mathbb{Z}$ replaced by $\widehat{h}_\gamma$. Before delineating the proof of Theorem 2, we state the following lemma which is a preparation for the proof of Theorem 2.

Lemma S. 1. *Let*

$$Q_{0R_1} = \frac{1}{S_{K,1}} \sum_{k'=1}^{K} \sum_{i'=1}^{n_{k'}} \sum_{j'=1}^{m_{k'i'}} W_{\widetilde{\eta}_{\mu,k'}}(T_{k'i'j'} - T_{kil}),$$

$$\Delta \widetilde{R}_1 = -S_{K,1}^{-1} S_{K,2}^{-1} \sum_{k=1}^{K} \sum_{i=1}^{n_k} \sum_{j \neq l} \sum_{k'=1}^{K} \sum_{i'=1}^{n_{k'}} \sum_{j'=1}^{m_{k'i'}} W_{\widetilde{\eta}_{\gamma,k'}}(T_{kij} - s) W_{\widetilde{\eta}_{\gamma,k'}}(T_{kil} - t) W_{\widetilde{\eta}_{\mu,k'}}(T_{k'i'j'} - T_{kil}),$$
$$\cdot (Y_{kij} - \mu(T_{kij})) \left\{ Y_{k'i'j'} - \mu(T_{k'i'j'}) - \frac{1}{2}\mu''(T_{kil})(T_{k'i'j'} - T_{kil})^2 \right\},$$

then

$$\mathrm{E}(\Delta \widetilde{R}_1 / Q_{0R_1}) = O_p(S_{K,1}^{-1} S_{K,2}^{-1} S_{K,3}),$$
$$\mathrm{Var}(\Delta \widetilde{R}_1 / Q_{0R_1}) = O_p\Big(S_{K,1}^{-2} S_{K,2}^{-2} \Big\{ S_{K,3}^2 + \sum_k s_{k,3} \widetilde{\eta}_{\mu,k}^{-1} \widetilde{\eta}_{\gamma,k}^{-2} + \sum_k s_{k,4} \widetilde{\eta}_{\mu,k}^{-1} \widetilde{\eta}_{\gamma,k}^{-1}$$
$$+ \sum_k s_{k,5} \widetilde{\eta}_{\gamma,k}^{-1} \Big\} + S_{K,2}^{-2} \sum_k s_{k,4} \widetilde{\eta}_{\gamma,k}^{4} \Big).$$

Proof. By easy calculation,

$$\mathrm{E}\Delta \widetilde{R}_1 = -S_{K,1}^{-1} S_{K,2}^{-1} \sum_{k=1}^{K} \sum_{i=1}^{n_k} \sum_{j \neq l} \sum_{k'=1}^{K} \sum_{i'=1}^{n_{k'}} \sum_{j'=1}^{m_{k'i'}} \mathrm{E}\big\{ W_{\widetilde{\eta}_{\gamma,k'}}(T_{kij} - s) W_{\widetilde{\eta}_{\gamma,k'}}(T_{kil} - t)$$
$$\cdot W_{\widetilde{\eta}_{\mu,k'}}(T_{k'i'j'} - T_{kil}) (Y_{kij} - \mu(T_{kij}))(Y_{k'i'j'} - \mu(T_{k'i'j'})) \big\}$$
$$= -S_{K,1}^{-1} S_{K,2}^{-1} \sum_{k=1}^{K} \sum_{i=1}^{n_k} \sum_{j \neq l} \sum_{i'=1}^{m_{ki}} \mathrm{E}\big\{ W_{\widetilde{\eta}_{\gamma,k}}(T_{kij} - s) W_{\widetilde{\eta}_{\gamma,k}}(T_{kil} - t) W_{\widetilde{\eta}_{\mu,k}}(T_{kij'} - T_{kil})$$
$$\cdot (Y_{kij} - \mu(T_{kij}))(Y_{kij'} - \mu(T_{kij'})) \big\}$$
$$= -S_{K,1}^{-1} S_{K,2}^{-1} \sum_{k=1}^{K} \Bigg[\sum_j m_{ki}^2 (m_{ki} - 1) \mathrm{E}\big\{ W_{\widetilde{\eta}_{\gamma,k}}(T_{kij} - s) W_{\widetilde{\eta}_{\gamma,k}}(T_{kil} - t)$$

$$
\left. \cdot W_{\widetilde{\eta}_{\mu,k}}(T_{kij'} - T_{kil})(Y_{kij} - \mu(T_{kij}))(Y_{kij'} - \mu(T_{kij'}))\} \right]
$$

$$
= O_p\left(S_{K,1}^{-1}S_{K,2}^{-1}S_{K,3}\right),
$$

and hence $\mathrm{E}\Delta\widetilde{R}_1/Q_{0R_1} = O_p(S_{K,1}^{-1}S_{K,2}^{-1}S_{K,3})$. Note that

$$
\begin{aligned}
\Delta\widetilde{R}_1^2 =& S_{K,1}^{-2}S_{K,2}^{-2}\sum_{j_1\neq l_1, j_2\neq l_2} W_{k_1i_1j_1}(s)W_{k_1i_1l_1}(t)W_{k_1'i_1'j_1'}(T_{k_1i_1l_1})W_{k_2i_2j_2}(s)W_{k_2i_2l_2}(t)W_{k_2'i_2'j_2'}(T_{k_2i_2l_2}) \\
&\cdot (Y_{k_1i_1j_1} - \mu_{k_1i_1j_1})\left\{Y_{k_1'i_1'j_1'} - \mu_{k_1'i_1'j_1'} - \frac{1}{2}\mu''_{k_1i_1l_1}(T_{k_1'i_1'j_1'} - T_{k_1i_1l_1})^2\right\} \\
&\cdot (Y_{k_2i_2j_2} - \mu_{k_2i_2j_2})\left\{Y_{k_2'i_2'j_2'} - \mu_{k_2'i_2'j_2'} - \frac{1}{2}\mu''_{k_2i_2l_2}(T_{k_2'i_2'j_2'} - T_{k_2i_2l_2})^2\right\},
\end{aligned}
$$

$\mathrm{E}\Delta\widetilde{R}_1^2$ is the summation of the following quantities:

$$
\begin{aligned}
(a) =& S_{K,1}^{-2}S_{K,2}^{-2}\mathrm{E}\sum W_{k_1i_1j_1}(s)W_{k_1i_1l_1}(t)W_{k_1'i_1'j_1'}(T_{k_1i_1l_1})W_{k_2i_2j_2}(s)W_{k_2i_2l_2}(t)W_{k_2'i_2'j_2'}(T_{k_2i_2l_2}) \\
&\cdot (Y-\mu)_{k_1i_1j_1}(Y-\mu)_{k_1'i_1'j_1'}(Y-\mu)_{k_2i_2j_2}(Y-\mu)_{k_2'i_2'j_2'}, \\
(b) =& -\frac{1}{2}S_{K,1}^{-2}S_{K,2}^{-2}\mathrm{E}\sum W_{k_1i_1j_1}(s)W_{k_1i_1l_1}(t)W_{k_1'i_1'j_1'}(T_{k_1i_1l_1})W_{k_2i_2j_2}(s)W_{k_2i_2l_2}(t)W_{k_2'i_2'j_2'}(T_{k_2i_2l_2}) \\
&\cdot (Y-\mu)_{k_1i_1j_1}(Y-\mu)_{k_1'i_1'j_1'}(Y-\mu)_{k_2i_2j_2}\mu''_{k_2i_2l_2}(T_{k_2'i_2'j_2'} - T_{k_2i_2l_2})^2, \\
(c) =& -\frac{1}{2}S_{K,1}^{-2}S_{K,2}^{-2}\mathrm{E}\sum W_{k_1i_1j_1}(s)W_{k_1i_1l_1}(t)W_{k_1'i_1'j_1'}(T_{k_1i_1l_1})W_{k_2i_2j_2}(s)W_{k_2i_2l_2}(t)W_{k_2'i_2'j_2'}(T_{k_2i_2l_2}) \\
&\cdot (Y-\mu)_{k_1i_1j_1}(Y-\mu)_{k_2i_2j_2}(Y-\mu)_{k_2'i_2'j_2'}\mu''_{k_1i_1l_1}(T_{k_1'i_1'j_1'} - T_{k_1i_1l_1})^2, \\
(d) =& \frac{1}{4}S_{K,1}^{-2}S_{K,2}^{-2}\mathrm{E}\sum W_{k_1i_1j_1}(s)W_{k_1i_1l_1}(t)W_{k_1'i_1'j_1'}(T_{k_1i_1l_1})W_{k_2i_2j_2}(s)W_{k_2i_2l_2}(t)W_{k_2'i_2'j_2'}(T_{k_2i_2l_2}) \\
&\cdot (Y-\mu)_{k_1i_1j_1}(Y-\mu)_{k_2i_2j_2}\mu''_{k_1i_1l_1}(T_{k_1'i_1'j_1'} - T_{k_1i_1l_1})^2\mu''_{k_2i_2l_2}(T_{k_2'i_2'j_2'} - T_{k_2i_2l_2})^2,
\end{aligned}
$$

which are computed as follows.

a. For (a), there are four cases according to the relationship between (k_1, i_1), (k_2, i_2), (k_1', i_1'), (k_2', i_2') to make the quantity nonzero:

(B.1) Let $\Lambda_1^a = \{(k_1, i_1) = (k_1', i_1'), (k_2, i_2) = (k_2', i_2'), (k_1, i_1) \neq (k_2, i_2), j_1 \neq l_1, j_2 \neq l_2\}$,

$$
\begin{aligned}
(a.1) =& S_{K,1}^{-2}S_{K,2}^{-2}\mathrm{E}\sum_{\Lambda_1^a} W_{k_1i_1j_1}(s)W_{k_1i_1l_1}(t)W_{k_1i_1'j_1}(T_{k_1i_1l_1})W_{k_2i_2j_2}(s)W_{k_2i_2l_2}(t) \\
&\cdot W_{k_2i_2j_2'}(T_{k_2i_2l_2})(Y-\mu)_{k_1i_1j_1}(Y-\mu)_{k_1i_1'j_1}(Y-\mu)_{k_2i_2j_2}(Y-\mu)_{k_2i_2j_2'} \\
=& S_{K,1}^{-2}S_{K,2}^{-2}\sum_{\Lambda_1^a}\mathrm{E}\left\{W_{k_1i_1l_1}(t)W_{k_1i_1j_1}(s)W_{k_1i_1'j_1}(T_{k_1i_1l_1})(Y-\mu)_{k_1i_1j_1}(Y-\mu)_{k_1i_1'j_1}\right\}
\end{aligned}
$$

$$\cdot \mathrm{E}\left\{W_{k_2i_2l_2}(t)W_{k_2i_2j_2}(s)W_{k_2i_2j_2'}(T_{k_2i_2l_2})(Y-\mu)_{k_2i_2j_2}(Y-\mu)_{k_2i_2j_2'}\right\}$$

$$=(\mathrm{E}\Delta\widetilde{R}_1(s,t))^2.$$

(B.2) Let $\Lambda_2^a = \{(k_1,i_1)=(k_2,i_2),(k_1',i_1')=(k_2',i_2'),(k_1,i_1)\neq(k_1',i_1'),j_1\neq l_1,j_2\neq l_2,j_1'\neq j_2'\}$,

$$\begin{aligned}(a.2) =& S_{K,1}^{-2}S_{K,2}^{-2}\sum_{\Lambda_2^a}\mathrm{E}\{W_{k_1i_1j_1}(s)W_{i_2j_1k_1}(s)(Y-\mu)_{k_1i_1j_1}(Y-\mu)_{i_2j_1k_1}\}\\ &\cdot \mathrm{E}\left[W_{k_1i_1l_1}(t)W_{l_2j_1k_1}(t)W_{k_1'i_1'j_1'}(T_{k_1i_1l_1})W_{i_2'j_1'k_1'}(T_{l_2j_1k_1})(Y-\mu)_{k_1'i_1'j_1'}(Y-\mu)_{i_2'j_1'k_1'}\right]\\ =& S_{K,1}^{-2}S_{K,2}^{-2}\sum_{\Lambda_2^a}\gamma(s,s)\gamma(t,t)f^2(s)f^4(t)+O_p(\rho_{\gamma,2}+\rho_{\mu,2})\\ =& S_{K,1}^{-2}S_{K,2}^{-2}S_{K,3}^2\left\{\gamma(s,s)\gamma(t,t)f^2(s)f^4(t)+O_p(\rho_{\gamma,2}+\rho_{\mu,2})\right\}.\end{aligned}$$

(B.3) Let $\Lambda_3^a = \{(k_1,i_1)=(k_2',i_2'),(k_2,i_2)=(k_1',i_1'),(k_1,i_1)\neq(k_2,i_2),j_1\neq l_1,j_2\neq l_2\}$,

$$\begin{aligned}(a.3) =& S_{K,1}^{-2}S_{K,2}^{-2}\mathrm{E}\Big[\sum_{\Lambda_3^a}W_{k_1i_1j_1}(s)W_{k_1i_1l_1}(t)W_{k_2i_2j_1'}(T_{k_1i_1l_1})\\ &\cdot W_{k_2i_2j_2}(s)W_{k_2i_2l_2}(t)W_{k_1i_1j_2'}(T_{k_2i_2l_2})\\ &\cdot (Y-\mu)_{k_1i_1j_1}(Y-\mu)_{k_2i_2j_1'}(Y-\mu)_{k_2i_2j_2}(Y-\mu)_{k_1i_1j_2'}\Big]\\ =& (\mathrm{E}\Delta\widetilde{R}_1(s,t))^2.\end{aligned}$$

(B.4) Let $\Lambda_4^a = \{(k_1,i_1)=(k_2,i_2)=(k_1',i_1')=(k_2',i_2'),j_1\neq l_1,j_2\neq l_2\}$ and omit the subscript ki,

$$\begin{aligned}(a.4) =& S_{K,1}^{-2}S_{K,2}^{-2}\mathrm{E}\Big\{\sum_{\Lambda_4^a}W_{j_1}(s)W_{l_1}(t)W_{j_1'}(T_{l_1})W_{j_2}(s)W_{l_2}(t)W_{j_2'}(T_{l_2})\\ &\cdot (Y-\mu)_{j_1}(Y-\mu)_{j_1'}(Y-\mu)_{j_2}(Y-\mu)_{j_2'}\Big\}.\end{aligned}$$

It is the summation of the following sets,

$\Lambda_{4.1}^a = \{j_1=j_2,j_1'=j_2',l_1=l_2,j_1\neq l_1,j_2\neq l_2\}$,

$\Lambda_{4.2}^a = \{j_1=j_2,j_1'\neq j_2',j_1\neq l_1,j_2\neq l_2\}\cup\{j_1\neq j_2,j_1'=j_2',l_1=l_2,j_1\neq l_1,j_2\neq l_2\}$,

$\Lambda_{4.3}^a = \{j_1,j_2,j_1',j_2'$ are not equal, $j_1\neq l_1,j_2\neq l_2\}$.

Correspondingly, we define

$$
\begin{aligned}
E_{1|4}(s,t) &= \mathrm{E}\big\{\big(Y_1-\mu(T_1)\big)^2\big(Y_2-\mu(T_2)\big)^2|T_1=s,T_2=t\big\},\\
E_{2|4}(s,t) &= \mathrm{E}\big\{\big(Y_1-\mu(T_1)\big)^2\big(Y_2-\mu(T_2)\big)\big)\big(Y_3-\mu(T_3)\big)|T_1=s,T_2=t,T_3=t\big\},\\
E_{3|4}(s,t) &= \mathrm{E}\big\{\big(Y_1-\mu(T_1)\big)\big(Y_2-\mu(T_2)\big)\big(Y_3-\mu(T_3)\big)\big(Y_4-\mu(T_4)\big)|T_1=s,\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad T_2=t,T_3=s,T_4=t\big\}.
\end{aligned}
$$

Then by easy calculation,

$$
\begin{aligned}
(a.4)\big|_{\Lambda^a_{4.1}} &\doteq S_{K,1}^{-2}S_{K,2}^{-2}\left(\sum_{k=1}^{K}s_{k,3}\widetilde{\eta}_{\mu,k}^{-1}\widetilde{\eta}_{\gamma,k}^{-2}\right)R(W)^3E_{1|4}(s,t)f(s)f^2(t),\\
(a.4)\big|_{\Lambda^a_{4.2}} &\doteq S_{K,1}^{-2}S_{K,2}^{-2}\Bigg\{\left(\sum_{k=1}^{K}s_{k,5}\widetilde{\eta}_{\gamma,k}^{-1}\right)R(W)E_{2|4}(s,t)f(s)f^4(t)\\
&\qquad+\left(\sum_{k=1}^{K}s_{k,4}\widetilde{\eta}_{\mu,k}^{-1}\widetilde{\eta}_{\gamma,k}^{-1}\right)R(W)^2E_{2|4}(t,s)f^2(s)f^2(t)\Bigg\},\\
(a.4)\big|_{\Lambda^a_{4.3}} &\doteq S_{K,1}^{-2}S_{K,2}^{-2}S_6E_{3|4}(s,t)f^2(s)f^4(t).
\end{aligned}
$$

b. For (b) and (c), it is required that $(k_1,i_1)=(k_2,i_2)=(k_1',i_1')$, define

$$
E_{2|3}(s,s,t)=\mathrm{E}\big\{\big(Y_1-\mu(T_1)\big)\big(Y_2-\mu(T_2)\big)\big(Y_3-\mu(T_3)\big)|T_1=s,T_2=s,T_3=t\big\}.
$$

By easy calculation,

$$
(b)+(c)\doteq S_{K,1}^{-2}S_{K,2}^{-2}\alpha(W)E_{2|3}(s,s,t)\left(\sum_{k=1}^{K}s_{k,3}\widetilde{\eta}_{\mu,k}^{2}\right).
$$

c. For (d), we need $(k_1,i_1)=(k_2,i_2)$, then

$$
(d)\doteq\frac{1}{4}S_{K,2}^{-2}\alpha^2(W)\{\mu''(t)\}^2\gamma(s,s)f^2(s)f^4(t)\left(\sum_{k=1}^{K}s_{k,4}\widetilde{\eta}_{\gamma,k}^{4}\right).
$$

From the above arguments,

$$
\begin{aligned}
\operatorname{Var}\Delta\widetilde{R}_1 &= \mathrm{E}\Delta\widetilde{R}_1^2-(\mathrm{E}\Delta\widetilde{R}_1)^2\\
&= O_p\Big(S_{K,1}^{-2}S_{K,2}^{-2}\Big\{S_{K,3}^2+\sum_k s_{k,3}\widetilde{\eta}_{\mu,k}^{-1}\widetilde{\eta}_{\gamma,k}^{-2}+\sum_k s_{k,4}\widetilde{\eta}_{\mu,k}^{-1}\widetilde{\eta}_{\gamma,k}^{-1}+\sum_k s_{k,5}\widetilde{\eta}_{\gamma,k}^{-1}\Big\}\\
&\qquad\qquad+S_{K,2}^{-2}\sum_k s_{k,4}\widetilde{\eta}_{\gamma,k}^{4}\Big).
\end{aligned}
$$

Note that

$$
\begin{aligned}
\mathrm{Cov}(Q_{0R_1}, \Delta\widetilde{R}_1) &= -S_{K,1}^{-2} S_{K,2}^{-1} \left(\sum_{k=1}^{K} \sum_{i=1}^{n_k} m_{ki}^2 (m_{ki} - 1) \right) \rho_{\mu,-1} R(W) \gamma(s,t) f(s) f^3(t), \\
\mathrm{E}(Q_{0R_1}) &= f(t) + \frac{1}{2} \alpha(W) f''(t) \rho_{\mu,2} + o_p\left(\rho_{\mu,2}\right), \\
\mathrm{Var}(Q_{0R_1}) &= \left(\frac{\rho_{\mu,-1}}{S_{K,1}} \right) R(W) f(t) + o_p \left(\frac{\rho_{\mu,-1}}{S_{K,1}} \right).
\end{aligned}
$$

By the delta method,

$$
\begin{aligned}
\mathrm{Var}\left(\frac{\Delta\widetilde{R}_1}{Q_{0R_1}} \right) &= \frac{\mathrm{Var}\,\Delta\widetilde{R}_1}{\mathrm{E}^2 Q_{0R_1}} - 2 \frac{\mathrm{E}\Delta\widetilde{R}_1 \,\mathrm{Cov}(Q_{0R_1}, \Delta\widetilde{R}_1)}{\mathrm{E}^3 Q_{0R_1}} + \frac{\mathrm{E}^2 \Delta\widetilde{R}_1 \,\mathrm{Var}\, Q_{0R_1}}{\mathrm{E}^4 Q_{0R_1}} \\
&= O_p\Big(S_{K,1}^{-2} S_{K,2}^{-2} \Big\{ S_{K,3}^2 + \sum_k s_{k,3} \widetilde{\eta}_{\mu,k}^{-1} \widetilde{\eta}_{\gamma,k}^{-2} + \sum_k s_{k,4} \widetilde{\eta}_{\mu,k}^{-1} \widetilde{\eta}_{\gamma,k}^{-1} + \sum_k s_{k,5} \widetilde{\eta}_{\gamma,k}^{-1} \Big\} \\
&\qquad + S_{K,2}^{-2} \sum_k s_{k,4} \widetilde{\eta}_{\gamma,k}^{4} \Big).
\end{aligned}
$$

■

Now we give the proof of the asymptotic distribution of $\widetilde{\gamma}^{(K)}$ in Theorem 2.

Proof. Denote

$$
f_1 = Q_{20} Q_{02} - Q_{11}^2, \quad f_2 = Q_{10} Q_{02} - Q_{01} Q_{11}, \quad f_3 = Q_{01} Q_{20} - Q_{10} Q_{11}.
$$

Let $\widetilde{\gamma}$ be the local linear estimate based on $\widetilde{C}_{kijl} = (Y_{kij} - \widetilde{\mu}_{kij})(Y_{kil} - \widetilde{\mu}_{kil})$ which can be expressed as

$$
\widetilde{\gamma} = \frac{f_1 R_{00} - f_2 R_{10} - f_3 R_{01}}{f_1 Q_{00} - f_2 Q_{10} - f_3 Q_{01}},
$$

where for $p, q = 0, 1, 2$,

$$
\begin{aligned}
Q_{p,q} &= S_{K,2}^{-1} \sum_{k=1}^{K} \sum_{i=1}^{n_k} \sum_{1 \le j \ne l \le m_{ki}} W_{\widetilde{\eta}_{\gamma,k}}(T_{kij} - s) W_{\widetilde{\eta}_{\gamma,k}}(T_{kil} - t) \left(\frac{T_{kij} - s}{\widetilde{\eta}_{\gamma,k}} \right)^p \left(\frac{T_{kil} - t}{\widetilde{\eta}_{\gamma,k}} \right)^q \\
&= \sum_k \omega_k Q_{pq,k}, \\
R_{p,q} &= S_{K,2}^{-1} \sum_{k=1}^{K} \sum_{i=1}^{n_k} \sum_{1 \le j \ne l \le m_{ki}} W_{\widetilde{\eta}_{\gamma,k}}(T_{kij} - s) W_{\widetilde{\eta}_{\gamma,k}}(T_{kil} - t) \left(\frac{T_{kij} - s}{\widetilde{\eta}_{\gamma,k}} \right)^p \left(\frac{T_{kil} - t}{\widetilde{\eta}_{\gamma,k}} \right)^q \widetilde{C}_{kijl} \\
&= \sum_k \omega_k R_{pq,k}.
\end{aligned}
$$

Let

$$\widetilde{\beta}_{1,k} = \frac{1}{\widetilde{\eta}_{\gamma,k}} \frac{f_2(R_{10} - R_{00}Q_{10}/Q_{00})}{f_1Q_{10} - f_2Q_{10}^2/Q_{00} - f_3Q_{01}Q_{10}/Q_{00}},$$
$$\widetilde{\beta}_{2,k} = \frac{1}{\widetilde{\eta}_{\gamma,k}} \frac{f_3(R_{01} - R_{00}Q_{01}/Q_{00})}{f_1Q_{01} - f_2Q_{01}Q_{10}/Q_{00} - f_3Q_{01}^2/Q_{00}},$$
$$\widetilde{\beta}^*_{1,k} = \frac{1}{\widetilde{\eta}_{\gamma,k}} \frac{f_2(R^*_{10} - R^*_{00}Q_{10}/Q_{00})}{f_1Q_{10} - f_2Q_{10}^2/Q_{00} - f_3Q_{01}Q_{10}/Q_{00}},$$
$$\widetilde{\beta}^*_{2,k} = \frac{1}{\widetilde{\eta}_{\gamma,k}} \frac{f_3(R^*_{01} - R^*_{00}Q_{01}/Q_{00})}{f_1Q_{01} - f_2Q_{01}Q_{10}/Q_{00} - f_3Q_{01}^2/Q_{00}},$$
$$R^*_{p,q} = S_{K,2}^{-1} \sum_{k=1}^{K} \sum_{i=1}^{n_k} \sum_{1\le j\ne l\le m_{ki}} W_{\widetilde{\eta}_{\gamma,k}}(T_{kij} - s) W_{\widetilde{\eta}_{\gamma,k}}(T_{kil} - t) \left(\frac{T_{kij} - s}{\widetilde{\eta}_{\gamma,k}}\right)^p \left(\frac{T_{kil} - t}{\widetilde{\eta}_{\gamma,k}}\right)^q C_{kijl},$$
$$\widetilde{G}_{pq} = S_{K,2}^{-1} \sum_{k=1}^{K} \sum_{i=1}^{n_k} \sum_{1\le j\ne l\le m_{ki}} W_{\widetilde{\eta}_{\gamma,k}}(T_{kij} - s) W_{\widetilde{\eta}_{\gamma,k}}(T_{kil} - t) T_{kij}^p T_{kil}^q = \sum_k \omega_k \widetilde{G}_{pq,k}.$$

Then we can write

$$\widetilde{\gamma} = \frac{R_{00}}{Q_{00}} - \sum_k \omega_k \widetilde{\beta}_{1,k} \cdot \frac{\widetilde{G}_{10,k} - sQ_{00}}{Q_{00}} - \sum_k \omega_k \widetilde{\beta}_{2,k} \cdot \frac{\widetilde{G}_{01,k} - tQ_{00}}{Q_{00}},$$
$$\overline{\gamma} = \frac{R_{00}}{Q_{00}} - \frac{\partial\gamma}{\partial s} \cdot \frac{\widetilde{G}_{10} - sQ_{00}}{Q_{00}} - \frac{\partial\gamma}{\partial t} \cdot \frac{\widetilde{G}_{01} - tQ_{00}}{Q_{00}},$$
$$\overline{\gamma}^* = \frac{R^*_{00}}{Q_{00}} - \frac{\partial\gamma}{\partial s} \cdot \frac{\widetilde{G}_{10} - sQ_{00}}{Q_{00}} - \frac{\partial\gamma}{\partial t} \cdot \frac{\widetilde{G}_{01} - tQ_{00}}{Q_{00}}.$$

Following similar arguments in the proof of Theorem 1, we can prove that

$$\widetilde{\gamma} - \overline{\gamma} = o_p\left[\min\left\{\left(S_{K,2}^{-2} \sum_{k=1}^{K} \sum_{i=1}^{n_k} \frac{s_{k,3}}{\widetilde{\eta}_{\gamma,k}}\right)^{\frac{1}{2}}, \left(S_{K,4}/S_{K,2}^2\right)^{\frac{1}{2}}\right\}\right].$$

Now we derive the asymptotic distribution of $\overline{\gamma}$. We first show that $\overline{\gamma} - \overline{\gamma}^* = O_p\{(N_K)^{-1}\}$.

(B.1) *Calculate $\overline{\gamma} - \overline{\gamma}^*$.*

From the definition of $\overline{\gamma}$ and $\overline{\gamma}^*$, we have $\overline{\gamma} - \overline{\gamma}^* = R_{00}/Q_{00}$. Note that

$$\begin{aligned} R_{00} =& S_{K,2}^{-1} \sum_{k=1}^{K} \sum_{i=1}^{n_k} \sum_{j\ne l} W_{kij}(s) W_{kil}(t) \widetilde{C}_{kijl} \\ =& S_{K,2}^{-1} \sum_{k=1}^{K} \sum_{i=1}^{n_k} \sum_{j\ne l} W_{kij}(s) W_{kil}(t) \{(Y_{kij} - \mu_{kij})(Y_{kil} - \mu_{kil}) + (Y_{kij} - \mu_{kij})(\mu_{kil} - \widetilde{\mu}_{kil}) \\ & + (\mu_{kij} - \widetilde{\mu}_{kil})(Y_{kil} - \mu_{kil}) + (\widetilde{\mu}_{kij} - \mu_{kij})(\widetilde{\mu}_{kil} - \mu_{kil})\} \end{aligned}$$

$$\doteq R_{00}^{*}+S_{K,2}^{-1}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\sum_{j\neq l}W_{kij}(s)W_{kil}(t)\{(Y_{kij}-\mu_{kij})(\mu_{kil}-\widetilde{\mu}_{kil})$$
$$+(\mu_{kij}-\widetilde{\mu}_{kil})(Y_{kil}-\mu_{kil})\}.$$

Recall that in S.1, we have

$$\widetilde{\mu}(t)-\overline{\mu}(t)=o_p(\overline{\mu}(t)-\mu(t)),$$

where

$$\begin{aligned}
\overline{\mu}(t)&=\frac{R_0}{Q_0}-\frac{\widetilde{G}_1}{Q_0}\mu'(t)+t\frac{\widetilde{G}_0}{Q_0}\mu'(t),\\
Q_r(t)&=\frac{1}{S_{K,1}}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\sum_{j=1}^{m_{ki}}W_{\widetilde{\eta}_{\mu,k}}(T_{kij}-t)\left(\frac{T_{kij}-t}{\widetilde{\eta}_{\mu,k}}\right)^{r},\\
R_r(t)&=\frac{1}{S_{K,1}}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\sum_{j=1}^{m_{ki}}W_{\widetilde{\eta}_{\mu,k}}(T_{kij}-t)\left(\frac{T_{kij}-t}{\widetilde{\eta}_{\mu,k}}\right)^{r}Y_{kij},\\
\widetilde{G}_r(t)&=\frac{1}{S_{K,1}}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\sum_{j=1}^{m_{ki}}W_{\widetilde{\eta}_{\mu,k}}(T_{kij}-t)T_{kij}^{r}.
\end{aligned}$$

then one can obtain

$$\begin{aligned}
&S_{K,2}^{-1}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\sum_{j\neq l}W_{kij}(s)W_{kil}(t)(Y_{kij}-\mu_{kij})(\mu_{kil}-\widetilde{\mu}_{kil})\\
\doteq&S_{K,2}^{-1}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\sum_{j\neq l}W_{kij}(s)W_{kil}(t)(Y_{kij}-\mu_{kij})(\mu_{kil}-\overline{\mu}_{kil})\\
=&S_{K,2}^{-1}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\sum_{j\neq l}W_{kij}(s)W_{kil}(t)(Y_{kij}-\mu_{kij})\left\{\mu(T_{kil})-\frac{R_0}{Q_0}+\frac{\check{G}_1}{Q_0}\mu'(T_{kil})-T_{kil}\frac{\check{G}_0}{Q_0}\mu'(T_{kil})\right\}\\
=&S_{K,2}^{-1}Q_0^{-1}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\sum_{j\neq l}W_{kij}(s)W_{kil}(t)(Y_{kij}-\mu_{kij})\\
&\cdot\left\{Q_0(T_{kil})\mu(T_{kil})-R_0(T_{kil})+\check{G}_1(T_{kil})\mu'(T_{kil})-T_{kil}\check{G}_0(T_{kil})\mu'(T_{kil})\right\}\\
=&S_{K,1}^{-1}S_{K,2}^{-1}Q_0^{-1}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\sum_{j\neq l}\sum_{k'=1}^{K}\sum_{i'=1}^{n_{k'}}\sum_{j'=1}^{m_{k'i'}}W_{\widetilde{\eta}_{\gamma,k'}}(T_{kij}-s)W_{\widetilde{\eta}_{\gamma,k'}}(T_{kil}-t)W_{\widetilde{\eta}_{\mu,k'}}(T_{k'i'j'}-T_{kil})\\
&\cdot(Y_{kij}-\mu(T_{kij}))\{\mu(T_{kil})-Y_{k'i'j'}+T_{k'i'j'}\mu'(T_{kil})-T_{kil}\mu'(T_{kil})\}.
\end{aligned}$$

Take the Taylor expansion of $\mu(T_{kil})$ at $\mu(T_{k'i'j'})$ in the last bracket,

$$S_{K,2}^{-1}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\sum_{j\neq l}W_{kij}(s)W_{kil}(t)(Y_{kij}-\mu_{kij})(\mu_{kil}-\widetilde{\mu}_{kil})$$

$$
\begin{aligned}
\doteq - S_{K,1}^{-1}S_{K,2}^{-1}Q_0^{-1}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\sum_{j\neq l}\sum_{k'=1}^{K}\sum_{i'=1}^{n_{k'}}\sum_{j'=1}^{m_{k'i'}} & W_{\widetilde{\eta}_{\gamma,k'}}(T_{kij}-s)W_{\widetilde{\eta}_{\gamma,k'}}(T_{kil}-t) \\
&\cdot W_{\widetilde{\eta}_{\mu,k'}}(T_{k'i'j'}-T_{kil})(Y_{kij}-\mu(T_{kij}))\Big\{Y_{k'i'j'}-\mu(T_{k'i'j'})+ \\
&\cdot \big(\mu'(T_{kil})-\mu'(T_{k'i'j'})\big)(T_{k'i'j'}-T_{kil})+\frac{1}{2}\mu''(T_{kil})(T_{k'i'j'}-T_{kil})^2\Big\}.
\end{aligned}
$$

Then take the Taylor expansion of $\mu'(T_{k'i'j'})$ at $\mu'(T_{kil})$ in the last bracket,

$$
\begin{aligned}
& S_{K,2}^{-1}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\sum_{j\neq l} W_{kij}(s)W_{kil}(t)(Y_{kij}-\mu_{kij})(\mu_{kil}-\widetilde{\mu}_{kil}) \\
\doteq & - S_{K,1}^{-1}S_{K,2}^{-1}Q_0^{-1}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\sum_{j\neq l}\sum_{k'=1}^{K}\sum_{i'=1}^{n_{k'}}\sum_{j'=1}^{m_{k'i'}} W_{\widetilde{\eta}_{\gamma,k'}}(T_{kij}-s)W_{\widetilde{\eta}_{\gamma,k'}}(T_{kil}-t) \\
& \cdot W_{\widetilde{\eta}_{\mu,k'}}(T_{k'i'j'}-T_{kil})(Y_{kij}-\mu(T_{kij}))\Big\{Y_{k'i'j'}-\mu(T_{k'i'j'})-\frac{1}{2}\mu''(T_{kil})(T_{k'i'j'}-T_{kil})^2\Big\}.
\end{aligned}
$$

Denote

$$
\begin{aligned}
\Delta\widetilde{R}_1 = & - S_{K,1}^{-1}S_{K,2}^{-1}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\sum_{j\neq l}\sum_{k'=1}^{K}\sum_{i'=1}^{n_{k'}}\sum_{j'=1}^{m_{k'i'}} W_{\widetilde{\eta}_{\gamma,k'}}(T_{kij}-s)W_{\widetilde{\eta}_{\gamma,k'}}(T_{kil}-t) \\
& \cdot W_{\widetilde{\eta}_{\mu,k'}}(T_{k'i'j'}-T_{kil})(Y_{kij}-\mu(T_{kij})\left\{Y_{k'i'j'}-\mu(T_{k'i'j'})-\frac{1}{2}\mu''(T_{kil})(T_{k'i'j'}-T_{kil})^2\right\}, \\
\Delta\widetilde{R}_2 = & - S_{K,1}^{-1}S_{K,2}^{-1}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\sum_{j\neq l}\sum_{k'=1}^{K}\sum_{i'=1}^{n_{k'}}\sum_{j'=1}^{m_{k'i'}} W_{\widetilde{\eta}_{\gamma,k'}}(T_{kij}-s)W_{\widetilde{\eta}_{\gamma,k'}}(T_{kil}-t) \\
& \cdot W_{\widetilde{\eta}_{\mu,k'}}(T_{k'i'j'}-T_{kij})(Y_{kil}-\mu(T_{kil})\left\{Y_{k'i'j'}-\mu(T_{k'i'j'})-\frac{1}{2}\mu''(T_{kij})(T_{k'i'j'}-T_{kij})^2\right\}, \\
Q_{0R1} = & Q_0(T_{kil}), \quad Q_{0R2} = Q_0(T_{kij}).
\end{aligned}
$$

It can be written as

$$
R_{00} \doteq R_{00}^{*} + \frac{\Delta\widetilde{R}_1}{Q_{0R_1}} + \frac{\Delta\widetilde{R}_2}{Q_{0R_2}},
$$

The last terms in the right hand side are of the same order. By Lemma S.1, we have

$$
\begin{aligned}
& \mathrm{E}(\Delta\widetilde{R}_1/Q_{0R_1}) = O_p(S_{K,1}^{-1}S_{K,2}^{-1}S_{K,3}), \\
& \mathrm{Var}(\Delta\widetilde{R}_1/Q_{0R_1}) = O_p\Big(S_{K,1}^{-2}S_{K,2}^{-2}\big\{S_{K,3}^2(1+S_{K,1}^{-1}\rho_{\mu,-1})+S_{K,3}\rho_{\mu,-1}\rho_{\gamma,-2} \\
& \qquad + S_4\rho_{\mu,-1}\rho_{\gamma,-1}+S_5\rho_{\gamma,-1}+S_6\big\}+S_{K,2}^{-2}S_4\left(\rho_{\mu,2}\right)^2\Big).
\end{aligned}
$$

Hence

$$
\begin{aligned}
\overline{\gamma}-\overline{\gamma}^{*} = & O_p\Big(S_{K,1}^{-1}S_{K,2}^{-1}\big\{S_{K,3}^2(1+S_{K,1}^{-1}\rho_{\mu,-1})+S_{K,3}\rho_{\mu,-1}\rho_{\gamma,-2} \\
& + S_{K,4}\rho_{\mu,-1}\rho_{\mu,-1}+S_5\rho_{\gamma,-1}+S_6\big\}^{\frac{1}{2}}+S_{K,2}^{-1}S_4^{\frac{1}{2}}\left(\rho_{\mu,2}\right)\Big).
\end{aligned}
$$

(B.2) *Asymptotic distribution of* $\overline{\gamma}^*$.

Define

$$\overline{\gamma}^* = \frac{R^*_{00}}{Q_{00}} - \frac{\partial\gamma}{\partial s}\cdot\frac{\widetilde{G}_{10} - sQ_{00}}{Q_{00}} - \frac{\partial\gamma}{\partial t}\cdot\frac{\widetilde{G}_{01} - tQ_{00}}{Q_{00}}.$$

The asymptotic normality of $\overline{\gamma}^*$ is accomplished by the asymptotic joint normality of $(R^*_{00} - \mathrm{E}R^*_{00}, \widetilde{G}_{00} - \mathrm{E}\widetilde{G}_{00}, \widetilde{G}_{10} - \mathrm{E}\widetilde{G}_{10}, \widetilde{G}_{01} - \mathrm{E}\widetilde{G}_{01})$. Specifically, let

$$\mathrm{E}_1(s,t) = \mathrm{E}\{(Y_1 - \mu(T_1))^2(Y_2 - \mu(T_2))^2 | T_1 = s, T_2 = t\},$$
$$\mathrm{E}_2(s,t) = \mathrm{E}\{(Y_1 - \mu(T_1))^2(Y_2 - \mu(T_2))(Y_2 - \mu(T_3)) | T_1 = s, T_2 = t, T_3 = t\}.$$

By easy calculation, we have

$$\begin{aligned}
\mathrm{E}\widetilde{G}_{pq} =& s^p t^q f(s)f(t) + \frac{1}{2}\alpha(W)\{s^p t^q f''(s)f(t) + s^p t^q f(s)f''(t)\}\cdot\overline{\eta^2_\gamma} \\
& + \frac{1}{2}\alpha(W)\{2pf'(s)f(t) + 2qf'(t)f(s)\}\cdot\overline{\eta^2_\gamma} + o_p\left(\overline{\eta^2_\gamma}\right), \\
\mathrm{E}R^*_{00} =& \gamma(s,t)f(s)f(t) + \frac{1}{2}\alpha(W)\overline{\eta^2_\gamma}\left\{\frac{\partial^2\gamma}{\partial s^2}(s,t)f(s)f(t)\right. \\
& + 2\frac{\partial\gamma}{\partial s}(s,t)f'(s)f(t) + \gamma(s,t)f''(s)f(t) + \frac{\partial^2\gamma}{\partial t^2}(s,t)f(t)f(s) \\
& \left. + 2\frac{\partial\gamma}{\partial}(s,t)f'(t)f(s) + \gamma(s,t)f''(t)f(s)\right\} + o_p\left(\overline{\eta^2_\gamma}\right), \\
\mathrm{Var}(R^*_{00}) =& \{1 + I(s=t)\}\left[R(W)^2 E_1(s,t)f(s)f(t)S^{-2}_{K,2}\overline{\eta^{-2}_\gamma}\right. \\
& \left. + R(W)\left\{f^2(s)f(t)E_2(t,s) + f(s)f(t)^2 E_2(s,t)\right\}S^{-1}_{K,2}\sum_{k=1}^K \frac{s_{k,3}}{\widetilde{\eta}_{\gamma,k}}\right] \\
& + V_3(s,t)f^2(s)f^2(t)S^{-2}_{K,2}S_4 + o\left(S^{-2}_{K,2}\overline{\eta^{-2}_\gamma}\right) \\
& + o_p\left(S^{-2}_{K,2}\sum_{k=1}^K \frac{s_{k,3}}{\widetilde{\eta}_{\gamma,k}}\right) + o_p\left(S^{-2}_{K,2}S_4\right), \\
\mathrm{Cov}(\widetilde{G}_{pq}, \widetilde{G}_{p'q'}) =& \{1 + I(s=t)\}s^{p+p'}t^{q+q'}\left[R(W)^2 f(s)f(t)S^{-2}_{K,2}\overline{\eta^{-2}_\gamma}\right. \\
& \left. + R(W)\{f^2(s)f(t) + f(s)f^2(t)\}S^{-2}_{K,2}\sum_{k=1}^K \frac{s_{k,3}}{\widetilde{\eta}_{\gamma,k}}\right] \\
& + o_p\left(S^{-2}_{K,2}\overline{\eta^{-2}_\gamma}\right) + o_p\left(S^{-2}_{K,2}\sum_{k=1}^K \frac{s_{k,3}}{\widetilde{\eta}_{\gamma,k}}\right), \\
\mathrm{Cov}(\widetilde{G}_{pq}, R^*_{00}) =& \{1 + I(s=t)\}s^p t^q\left[R(W)^2\gamma(s,t)f(s)f(t)S^{-2}_{K,2}\overline{\eta^{-2}_\gamma}\right.
\end{aligned}$$

$$+ R(W)\gamma(s,t)\{f^2(s)f(t) + f(s)f^2(t)\}S_{K,2}^{-2}\sum_{k=1}^{K}\frac{s_{k,3}}{\widetilde{\eta}_{\gamma,k}}\Bigg]$$

$$+ o_p\left(S_{K,2}^{-2}\overline{\eta_\gamma^{-2}}\right) + o_p\left(S_{K,2}^{-2}\sum_{k=1}^{K}\frac{s_{k,3}}{\widetilde{\eta}_{\gamma,k}}\right).$$

From the delta method, $\overline{\gamma}(t)$ follows the asymptotic normality that

$$\Gamma_{\overline{\gamma}^*}^{-1/2}\left\{\overline{\gamma}^*(s,t) - \gamma(s,t) - \frac{1}{2}\overline{\eta_\gamma^2}\alpha(W)\left(\frac{\partial^2\gamma}{\partial s^2}(s,t) + \frac{\partial^2\gamma}{\partial t^2}(s,t)\right) + o_p\left(\overline{\eta_\gamma^2}\right)\right\} \xrightarrow{d} N(0,1),$$

where

$$\Gamma_{\overline{\gamma}^*} = \{1 + I(s=t)\}\left\{S_{K,2}^{-1}\overline{\eta_\gamma^{-2}}R(W)^2\frac{V_1(s,t)}{f(s)f(t)}\right.$$
$$\left. + S_{K,2}^{-2}\sum_{k=1}^{K}\frac{s_{k,3}}{\widetilde{\eta}_{\gamma,k}}R(W)\frac{f(s)V_2(t,s) + f(t)V_2(s,t)}{f(s)f(t)}\right\} + V_3(s,t)S_4/S_{K,2}^2.$$

The proof is completed combining the above arguments. ■

S.3. PROOF OF THEOREM 3

We mention that Theorem 3 and Theorem 4 can be extended to the standard d-dimensional local linear regression. To see this, we introduce the following notations and assumptions.

Notations and assumptions for d-dimensional local linear regression. The underlying model is $Y = m(X) + \varepsilon$ where $X \in \mathcal{R}^d$, $\mathrm{E}\varepsilon = 0$ and $\mathrm{Var}\,\varepsilon = \sigma^2(X)$. Suppose we observe $\{(Y_{ki}, X_{ki} : i = 1, \dots, n_k)\}$ in the kth block and $N_K = \sum_{k=1}^K n_k$, i.e., we use n_K to denote the sub-sample size of the Kth data block and $N_K = \sum_{k=1}^K n_k$ to denote the full sample size up to time K, which correspond to $s_{K,1}, S_{K,1}$ in (6) for mean estimation and $s_{K,2}, S_{K,2}$ for covariance estimation, respectively. We impose the following assumptions which are in parallel to (A.1)–(A.3) and (A.6).

(B.1) Observations $\{X_{ki} : i = 1, \dots, n_k, k = 1, \dots, K\}$ are i.i.d. copies of a random variable X defined on $[0,1]^d$ whose density $f(\cdot)$ is bounded away from 0 with bounded second derivative.

(B.2) The second and fourth derivatives of regression function m(x) are bounded and continuous.

(B.3) Noises ε_{Ki} are i.i.d. with mean 0 and variance $\sigma^2(x)$.

(B.4) The block size $Kn_k/N_K \to 1$ as $K \to \infty$ for $k = 1, 2, \ldots, K$.

The optimal bandwidth at time K is $h_*^{(K)} \asymp N_K^{-1/(4+d)}$. The candidate sequence is $\{\eta_l^{(K)}\}_{l=1}^L$ and the pseudo-bandwidths are $\{\widetilde{\eta}_k^{(K)} : k = 1, \ldots, K\}$ for $j = 1, \ldots, d$. The centroids for the candidate sequence are $\{\phi_i^{(K)} : i = 1, \ldots, L\}$. Further define

$$\rho_i^{(K)} = \sum_{k=1}^{K} \frac{n_k}{N_K} \left(\widetilde{\eta}_k^{(K)}\right)^i .$$

We also mention that the mean and covariance estimates corresponds the cases $d = 1$ and $d = 2$, respectively.

When estimating μ, all observations are pooled together and the index of subjects become invalid, which is intrinsically one-dimensional local linear regression. Hence one can derive the following result for standard d-dimensional local linear regression by the same arguments in Theorem 1.

Lemma S. 2. *Suppose assumptions* (B.1)–(B.3) *and (A.8)–(A.9) hold. If*

$$\widetilde{h}^{(K)} - h_*^{(K)} = o_p\left(N_K^{-\frac{1}{4+d}}\right), \quad j = 1, \ldots, d, \tag{S.3}$$

then a fixed interior point $x \in (0,1)^d$, as $K \to \infty$, the online estimate $\widetilde{m}^{(K)}(x)$ satisfies

$$\left\{N_K/\rho_{-d}^{(K)}\right\}^{\frac{1}{2}} \left[\widetilde{m}^{(K)}(x) - m(x) - \frac{1}{2}\alpha(W)\left\{\sum_{j=1}^{d} m_j''(x)\right\}\rho_2 + o_p(\rho_2)\right] \xrightarrow{d} N\left(0, \frac{R(W)^d\sigma^2(x)}{f(x)}\right).$$

We prove in Section S.3 that (S.3) is satisfied by the proposed online method.

Denote $\mathbb{X} = \{X_{ki} \in \mathcal{R}^d : k = 1, \ldots, K; i = 1, \ldots, n_k\}$, $\theta = \int\{\sum_{j=1}^d \partial^2 m(x)/\partial x_j^2\} f(x)dx$ and $\nu = \int_{[0,1]^d} R(W)^d \sigma^2(x)dx$. Similarly, we adopt another online local cubic and online local linear to estimate θ and ν, which are denoted as $\widetilde{\theta}^{(K)}$ and $\widetilde{\nu}^{(K)}$, respectively. Then the optimal bandwidth and the online estimated bandwidth for $m(\cdot)$ are

$$h_*^{(K)} = \left(\frac{\nu}{\alpha^2(W)\theta}\right)^{\frac{1}{4+d}} N_K^{-\frac{1}{4+d}}, \quad \widetilde{h}^{(K)} = \left(\frac{\widetilde{\nu}^{(K)}}{\alpha^2(W)\widetilde{\theta}^{(K)}}\right)^{\frac{1}{4+d}} N_K^{-\frac{1}{4+d}}. \tag{S.4}$$

The candidate sequence for θ and ν are $\{\eta_{\theta,l}^{(K)}\}_{l=1}^J$ and $\{\eta_{\nu,l}^{(K)}\}_{l=1}^J$, respectively, and the pseudo-bandwidths are $\{\widetilde{\eta}_{\theta,k}^{(K)} : k = 1, \ldots, K\}$ and $\{\widetilde{\eta}_{\nu,k}^{(K)} : k = 1, \ldots, K\}$. Then we have the following conclusion.

Lemma S. 3. *Suppose assumptions* (B.1)–(B.4) *and (A.8)–(A.9) hold. When the online bandwidths and candidate sequence for θ and ν are*

$$h_\theta^{(K)} = GN_K^{-1/(6+d)}, \quad h_\nu^{(K)} = RN_K^{-1/(4+d)}, \quad 1 < G, R < \infty,$$
$$\eta_{\theta,J}^{(K)} < \cdots < \eta_{\theta,1}^{(K)} = h_\theta^{(K)}, \quad \eta_{\nu,J}^{(K)} < \cdots < \eta_{\nu,1}^{(K)} = h_\nu^{(K)},$$

then as $K \to \infty$, $\widetilde{h}^{(K)}$ in (S.4) *satisfies*

$$\frac{\widetilde{h}^{(K)} - h_*^{(K)}}{h_*^{(K)}} = O_p\left(N_K^{-\frac{2}{d+2}}\right).$$

Proof. Using the same technique in the proof S.1, we can derive that

$$\begin{aligned}\mathrm{E}\{(\widetilde{\theta}^{(K)} - \theta)^2 \mid \mathbb{X}\} &= \left(A_1\rho_{\theta,2} + A_2 N_K^{-1}\rho_{\theta,-4-d}\right)^2 + A_3 N_K^{-2}\rho_{\theta,-8-d} \\ &\quad + o_p\left(\left\{\rho_{\theta,2} + N_K^{-1}\rho_{\theta,-4-d}\right\}^2 + N_K^{-2}\rho_{\theta,-8-d}\right),\end{aligned}$$

where A_1, A_2, A_3 are the same constants as in Wand and Jones (1994), and

$$\rho_{\theta,i} = \frac{1}{N_K}\sum_{k=1}^{K} n_k(\widetilde{\eta}_{\theta,k})^i, \ i = 2, -5, -9.$$

By similar argument as S.4, the online estimates $\widetilde{\theta}^{(K)}$ satisfies

$$\frac{\widetilde{\theta}^{(K)} - \theta}{\widehat{\theta}^{(K)} - \theta} = O_p\left(1 + \frac{1}{J} + \frac{1}{J^2}\right),$$

i.e., $\widetilde{\theta}^{(K)} - \theta = O_p\big(N_K^{-2/(d+6)}\big)$, where J is the length of the candidate bandwidth sequence for estimating θ. Substituting the expression of $\widetilde{h}^{(K)}$ and $h_*^{(K)}$ in (S.4), we obtain

$$\frac{\widetilde{h}^{(K)} - h_*^{(K)}}{h_*^{(K)}} = \frac{1}{d+4}\left(\frac{\theta}{\widetilde{\theta}^{(K)}}\right)^{\frac{1}{d+4}}\left(\frac{\widetilde{\nu}^{(K)} - \nu}{\nu}\right) - \frac{1}{d+4}\frac{\widetilde{\theta}^{(K)} - \theta}{\theta}.$$

When $h_\nu^{(K)} \asymp N_K^{-1/5}$, $\widetilde{\sigma}^2(x) - \sigma^2(x) = O_p(N_K^{-2/5})$ and then $\widetilde{\nu} - \nu = O_p(N_K^{-2/5})$, and hence the convergence rate of $\widetilde{h}^{(K)}$ is dominated by $\widetilde{\theta}^{(K)} - \theta$. ∎

The online estimate $\widetilde{h}^{(K)}$ attains the same convergence rate as the batch estimate $\widehat{h}^{(K)}$. Similarly, in functional data analysis, one can prove that $\widetilde{\theta}_\mu^{(K)} - \theta_\mu = O_p(S_{K,1}^{-2/7})$ and $\widetilde{\theta}_\gamma^{(K)} - \theta_\gamma = O_p(S_{K,2}^{-1/4})$, and the convergence rate of $\widetilde{h}_\mu^{(K)}$ and $\widetilde{h}_\gamma^{(K)}$ are dominated by $\widetilde{\theta}_\mu^{(K)} - \theta_\mu$ and $\widetilde{\theta}_\gamma^{(K)} - \theta_\gamma$ respectively, in the four cases listed in Appendix B. Specifically, when estimating μ based on sparse data, from Fan and Yao (1998), when $h_\nu^{(K)} = O_p(S_{K,1}^{-1/5})$, $\widetilde{r}^{(K)}(t) - r(t) = O_p(S_{K,1}^{-2/5})$

and then $\widetilde{\nu}_\mu^{(K)} - \nu_\mu = O_p(S_{K,1}^{-2/5})$. For γ, one can prove that the convergence rate of $\widetilde{\nu}_\gamma^{(K)}$ is dominated by $\check{\mathrm{E}}_K \Phi(s)\Phi(t)$ in step 3 of case (b) in Appendix B which equals $O_p(S_{K,2}^{-1/3})$. For dense data, we need only check the order of $\widetilde{\sigma}_K^2$. From Zhang and Chen (2007), $\check{\varepsilon}_{kij} - \varepsilon_{kij}$ in step 2 of (c) in Appendix B is of order $(S_{K,1})^{-2/25}$, recall that $\widehat{\varepsilon}_{kij}^2$ is a debiased version of $\check{\varepsilon}_{kij}$ and $\widetilde{\sigma}_K^2$ is the average of $\widehat{\varepsilon}_{kij}^2$, then $\widetilde{\sigma}_K^2 - \sigma^2 = O_p\big(S_{K,1}^{-12/25}\big)$. In all the cases, estimates of θ are dominates those of ν, which completes the proof of Theorem 3.

S.4. PROOF OF THEOREM 4

Note that the asymptotic distributions of $\widetilde{\mu}$ and $\widetilde{\gamma}$ have the consistent form with the standard d-dimensional local linear regression, i.e., the bias is of $O_p(h^2)$ and the variance is of $O_p(N^{-1}h^{-d})$, where $d = 1$ for mean estimation and $d = 2$ for covariance estimation. The lower bound of the relative efficiency also holds for the standard d-dimensional local linear regression. We adopt the notations of the d-dimensional local linear in the proof for notation conciseness.

Proof. Recall that $\widetilde{h}^{(K)} = \eta_1^{(K)} > \ldots > \eta_L^{(K)}$, without loss of generality, we assume $\eta_l^{(K)} = g(l)\widetilde{h}^{(K)}$ where $g(l) \le g(1) = 1$, $l = 1, \ldots, L$. Note that $\widetilde{h}^{(K)} - \widehat{h}^{(K)} = o_p\big(N_K^{-1/(d+4)}\big)$ and the candidate bandwidth sequence satisfies

$$\frac{\widetilde{\eta}_k^{(K)}}{\widehat{h}^{(K)}} = g(l)\left(\frac{N_k}{N_K}\right)^{-\frac{1}{d+2}} + o_p(1).$$

Write $g(l) = \{g(l)^{1/\lambda}\}^\lambda$ and note that N_k/N_K grows linearly with respect to k, then λ shall be $1/(d+2)$ and the optimal $g(l)^{1/\lambda}$ shall be linear between $(0,1)$. Hence, the optimal η_l is as in (19).

Next we derive the asymptotic lower bound for the relative efficiency of our local linear estimator. Writing $\widehat{h}^{(K)} = h_*^{(K)}\{1 + O_p(N_K^{-2/(d+6)})\}$, one can derive

$$eff(\widetilde{m})^{-1} = \frac{d}{d+4}\left\{\sum_{k=1}^K \frac{n_k}{N_K}\left(\frac{\widetilde{\eta}_k^{(K)}}{\widetilde{h}^{(K)}}\right)^2\right\}^2 + \frac{4}{d+4}\left\{\sum_{k=1}^K \frac{n_k}{N_K}\left(\frac{\widetilde{\eta}_k^{(K)}}{\widetilde{h}^{(K)}}\right)^{-1}\right\} + O_p\left(N_K^{-\frac{2}{6+d}}\right). \tag{S.5}$$

As shown in Figure 1, when K tends large, there exists a breakpoint K_0: when $k \le K_0$, $\widetilde{\eta}_k$ equals the last candidate $(1/L)^{1/(d+4)}\widetilde{h}_k$ for the limit number of candidates and for $k > K_0$, there are sufficient and close candidates to make a choice. Hence

$$\frac{\widetilde{\eta}_k^{(K)}}{\widetilde{h}^{(K)}} = \begin{cases} L^{-\frac{1}{d+4}} \cdot \left(\frac{N_k}{N_K}\right)^{-\frac{1}{d+4}} + O_p\left(N_k^{-\frac{2}{6+d}}\right), & k \le K_0 \\ \left\{1 + \left(\frac{N_k}{N_K} - \frac{l_k}{L}\right)\right\}^{\frac{1}{d+4}} + O_p\left(N_k^{-\frac{2}{6+d}}\right), & k > K_0, \end{cases}$$

where $l_k = \mathrm{argmin}_{1\le l\le L} |N_k/N_K - l/L|$. Then when $k \le K_0$,

$$\sum_{k=1}^{K_0} \frac{n_k}{N_K}\left(\frac{\widetilde{\eta}_k^{(K)}}{\widetilde{h}^{(K)}}\right)^2 = \frac{d+4}{d+2}\cdot L^{-\frac{2}{d+4}}\left(\frac{N_{K_0}}{N_K}\right)^{1-\frac{2}{d+4}} + O_p\left(N_k^{-\frac{2}{6+d}}\right),$$
$$\sum_{k=1}^{K_0} \frac{n_k}{N_K}\left(\frac{\widetilde{\eta}_k^{(K)}}{\widetilde{h}^{(K)}}\right)^{-d} = \frac{d+4}{2d+4}\cdot L^{\frac{d}{d+4}}\left(\frac{N_{K_0}}{N_K}\right)^{1+\frac{d}{d+4}} + O_p\left(N_k^{-\frac{2}{6+d}}\right).$$

and when $k > K_0$,

$$\sum_{k=K_0+1}^{K} \frac{n_k}{N_K}\left(\frac{\widetilde{\eta}_k^{(K)}}{\widetilde{h}^{(K)}}\right)^2 = 1 - \frac{N_{K_0}}{N_K} + O_p\left(N_K^{-\frac{2}{6+d}}\right),$$
$$\sum_{k=K_0+1}^{K} \frac{n_k}{N_K}\left(\frac{\widetilde{\eta}_k^{(K)}}{\widetilde{h}^{(K)}}\right)^{-1} = 1 - \frac{N_{K_0}}{N_K} + O_p\left(N_K^{-\frac{2}{6+d}}\right).$$

The property of breakpoint K_0 also guarantees that

$$\left(\frac{1}{L}\right)^{\frac{1}{d+4}} \sum_{k=1}^{K_0} \frac{n_k}{N_{K_0}} \widetilde{h}^{(k)} \ge \widetilde{h}^{(K)},$$

which is equivalent to

$$\left(\frac{1}{L}\right)^{\frac{1}{d+4}} \sum_{k=1}^{K_0} \frac{n_k}{N_{K_0}} \frac{\widetilde{h}^{(k)}/h_*^{(k)}}{\widetilde{h}^{(K)}/h_*^{(K)}} \frac{h_*^{(k)}}{h_*^{(K)}} = \left(\frac{1}{L}\right)^{\frac{1}{d+4}} \sum_{k=1}^{K_0} \frac{n_k}{N_{K_0}} \{1 + O_p(N_K^{-\frac{2}{6+d}})\}\left(\frac{N_k}{N_K}\right)^{\frac{1}{d+4}} \ge 1,$$

and under Assumption (A.5), we obtain

$$\frac{N_{K_0}}{N_K} \le \left(\frac{d+4}{d+3}\right)^{d+4} \frac{1}{L} + O_p\left(N_K^{-\frac{2}{6+d}}\right).$$

Denote $\rho = N_{K_0}/N_K$, one can derive from (S.5) that

$$eff(\widetilde{m})^{-1} = \frac{d}{d+4}\left\{\frac{d+4}{d+2}\cdot L^{-\frac{2}{d+4}}\rho^{1-\frac{2}{d+4}} + 1 - \rho\right\}^2 + \frac{4}{d+4}\left\{\frac{d+4}{2d+4}\cdot L^{\frac{d}{d+4}}\rho^{1+\frac{d}{d+4}} + 1 - \rho\right\} + O_p\left(N_K^{-\frac{2}{6+d}}\right). \quad \text{(S.6)}$$

Note that (S.6) is strictly increasing with respect to $\rho \in \left[0, \{(d+4)/(d+3)\}^{d+4}/L\right]$. Hence we have

$$eff(\widetilde{m})^{-1} \leq 1 + \frac{2c_1 d + 4c_2}{d+4} L^{-1} + \frac{d}{d+4} c_1^2 L^{-2},$$

where

$$c_1 = \frac{(d+4)^{d+3}}{(d+3)^{d+2}(d+2)} - \left(\frac{d+4}{d+3}\right)^{d+4}, c_2 = \frac{(d+4)^{2d+5}}{(d+3)^{2d+4}(2d+4)} - \left(\frac{d+4}{d+3}\right)^{d+4}.$$

This completes the proof of Theorem 4.

■

REFERENCES

Fan, J. and Yao, Q. (1998), 'Efficient estimation of conditional variance functions in stochastic regression', *Biometrika* **85**(3), 645–660.

Wand, M. P. and Jones, M. C. (1994), *Kernel Smoothing*, Chapman and Hall.

Zhang, J. and Chen, J. (2007), 'Statistical inferences for functional data', *The Annals of Statistics* **35**(3), 1052–1079.

Zhang, X. and Wang, J.-L. (2016), 'From sparse to dense functional data and beyond', *The Annals of Statistics* **44**(5), 2281–2321.